Atmospheric Science Data Center Update

CERES Science Team Meeting

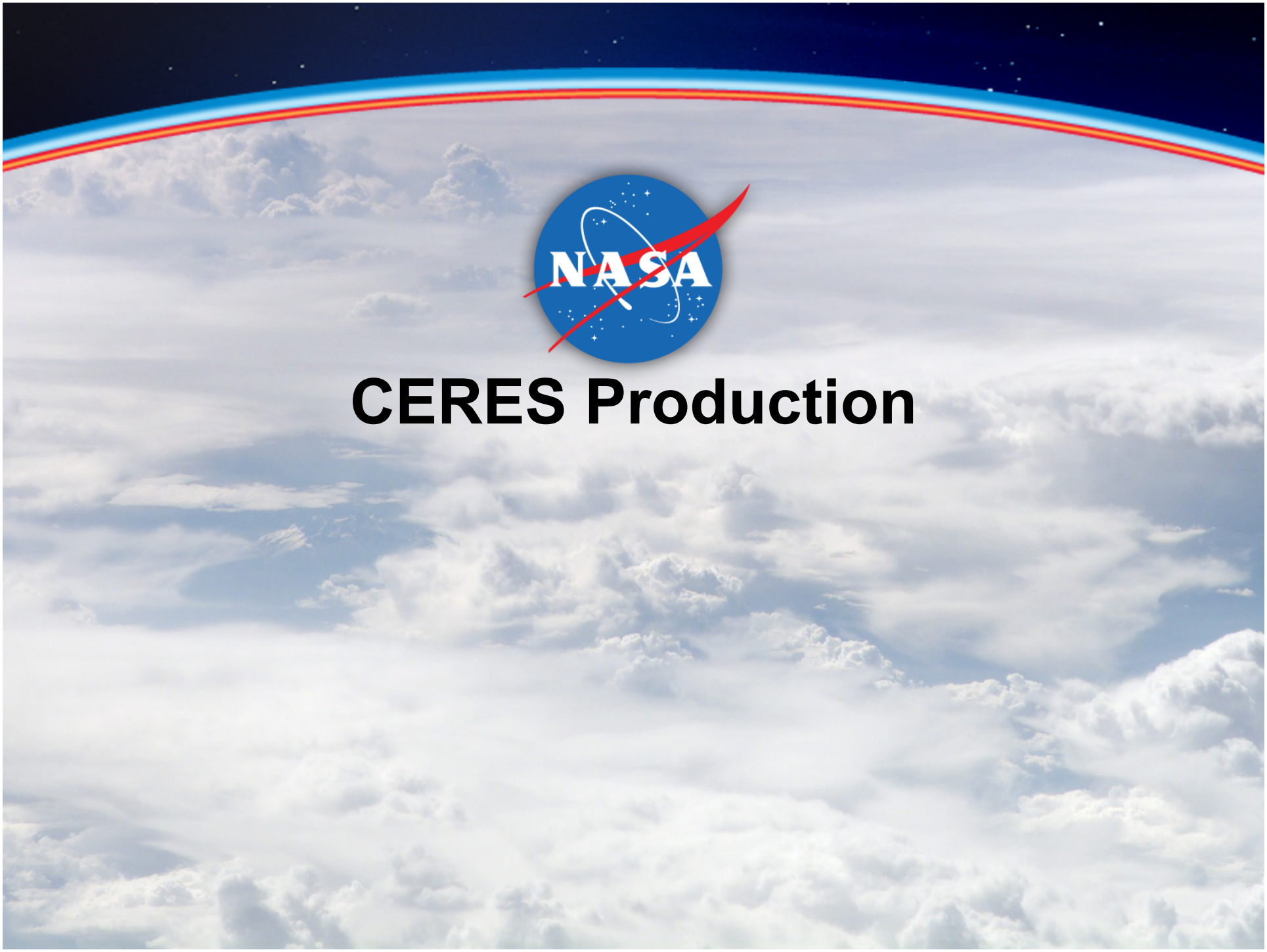
13 September 2010

Susan Sorlie, SSAI

Operations and User Services Manager

Topics

- CERES Production
  - AMI (ASDC Modernization through Integration)
- Distribution and Customer Metrics
- What's New at the ASDC
- Accessing CERES Data
- Contacting the ASDC User Services



CERES Production

CERES Production

- CERES processing environment is a blend of legacy and new systems
  - *Warlock*: SGI Origin 3800
  - *Magneto:* Linux Cluster with IBM PowerPC 970 Blades
  - *AMI*: IBM Linux system deployed to meet combined SCF and ASDC Production computing requirements; re-architecture is underway to convert the I/O subsystem from all-SAN to SAN/LAN topology to improve system stability and facilitate scaling beyond current size of ~ 250 servers
  - *AMI-P (AMI Precursor):* Linux Cluster with IBM Power6 and Intel x86 SMP servers to support SSI&T and production processing.
- Majority of science code already ported to Magneto and CERES Instrument code for NPP validated on AMI-P.

Ultimate goal is for all CERES production to be off legacy machines

CERES Production

CERES Production System	CPUs	Disk Space	Comments
Magneto *IBM Linux Cluster* *~5 years old*	112	39TB*	56 IBM JS20 P4 2-processor blades; requires constant supervision to keep CPUs up; DPO (Data Products On-line) NFS mounted over network for data staging; [*plan to increase production work space by 30 TB in Sept to total of 69 TB]
Warlock *SGI Origin 3800* *~10 years old*	128	32TB	Uses SGI RAIDs directly attached (CERES 18 TB; MODIS 14 TB; FLASH 7.5 TB);DPO NFS mounted over network for data staging; T&M maintenance started 9/1/2010
AMI-P *IBM Linux system*	112 @ P6 48 @ x86	99 TB	112 SGE P6 job slots, 48 Sun Grid Engine (SGE) x86 job slots; DPO mounted via fibre channel for rapid data access; made available for CERES Science SIT in mid-August 2010

CERES Production

- AMI Re-Architecture effort implemented in spring 2010 to address known issues and constraints with initial AMI design
  - Original AMI architecture was not expandable or stable
  - Performance requirements were not consistently met
  - Improvements had to fit within available funding
  - Configurations need to be proven and incorporate industry best practices
- AMI-P (Precursor) deployed as smaller stable system until AMI re-architecture is completed.

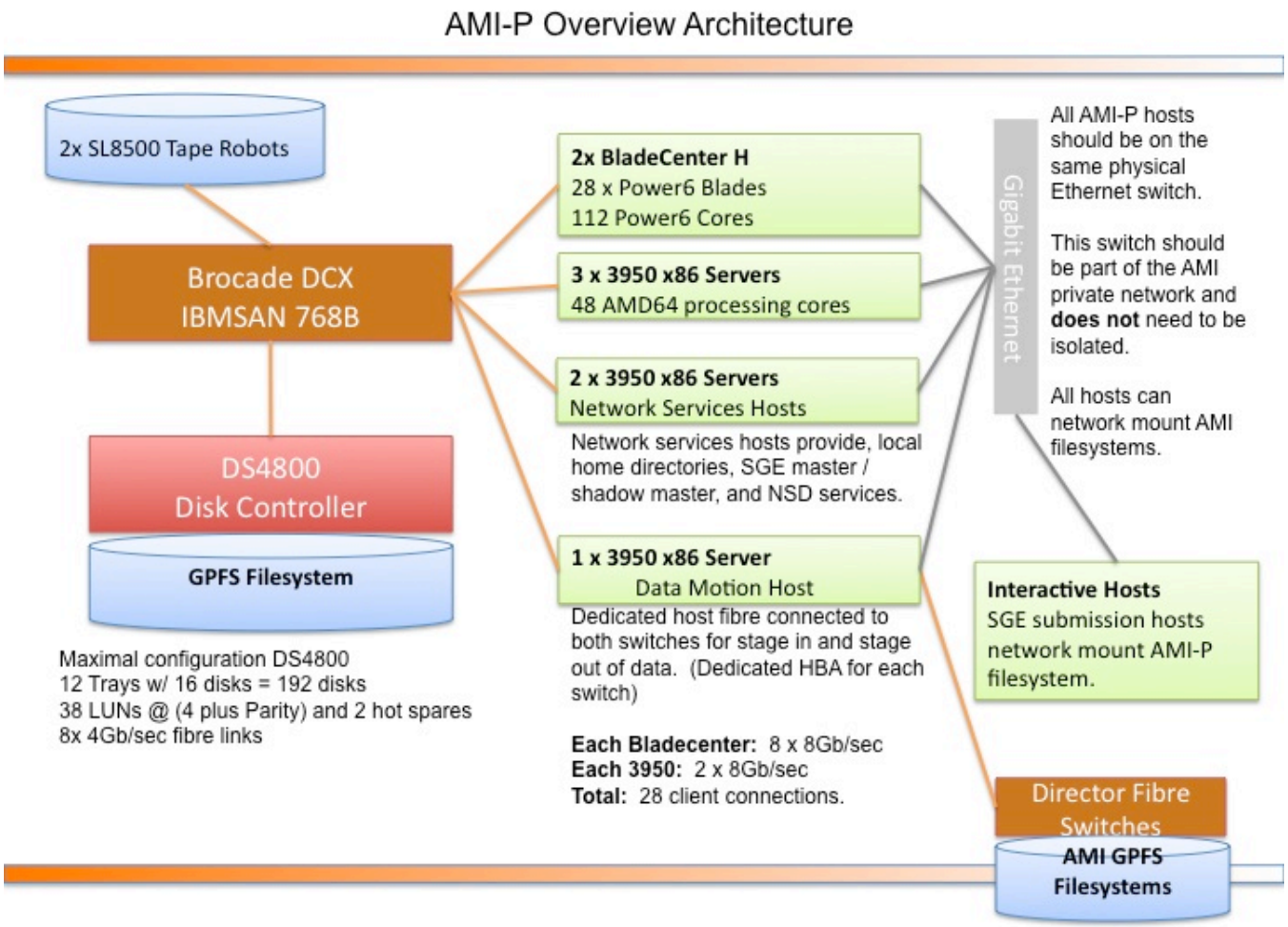
IBM Proof of Concept System

- As part of the AMI re-architecture work, the ASDC brought in the Proof of Concept (POC) system (valued at $1.5M) that was set up by IBM and DSS personnel in June and returned in mid-August .
  - Initially validated setup via multiple configurations with canned “acceptance test” results
  - It was used to determine performance on actual ASDC PGEs from CERES ,CALIPSO, and MISR. System activity performance data per blade was also captured for application to new hardware on order.
  - POC results helped to determine the procurement of test bed hardware for CLARREO

AMI-Precursor (AMI-P)

- AMI-P was certified as ready for production and released to CERES users late August.
  - AMI-P provides additional computational power to support Edition 2/3 and FM5 processing efforts.
  - Completed activities include: Installation and configuration of the Operations Environment, testing of Instrument Software Change Control Request (SCCR) 717, promotion of the following deliveries to AMI-P: CERESlib (SCCR 793), PERL lib (SCCR 780), Instrument Gains (Edition3), Instrument Gains (Edition2), Instrument Gains (Edition3 -FM1,FM2).
  - Coding and operational testing for CERES Instrument SCCR 716 is underway. This is the first delivery of the Ada software to the new X86 platform.

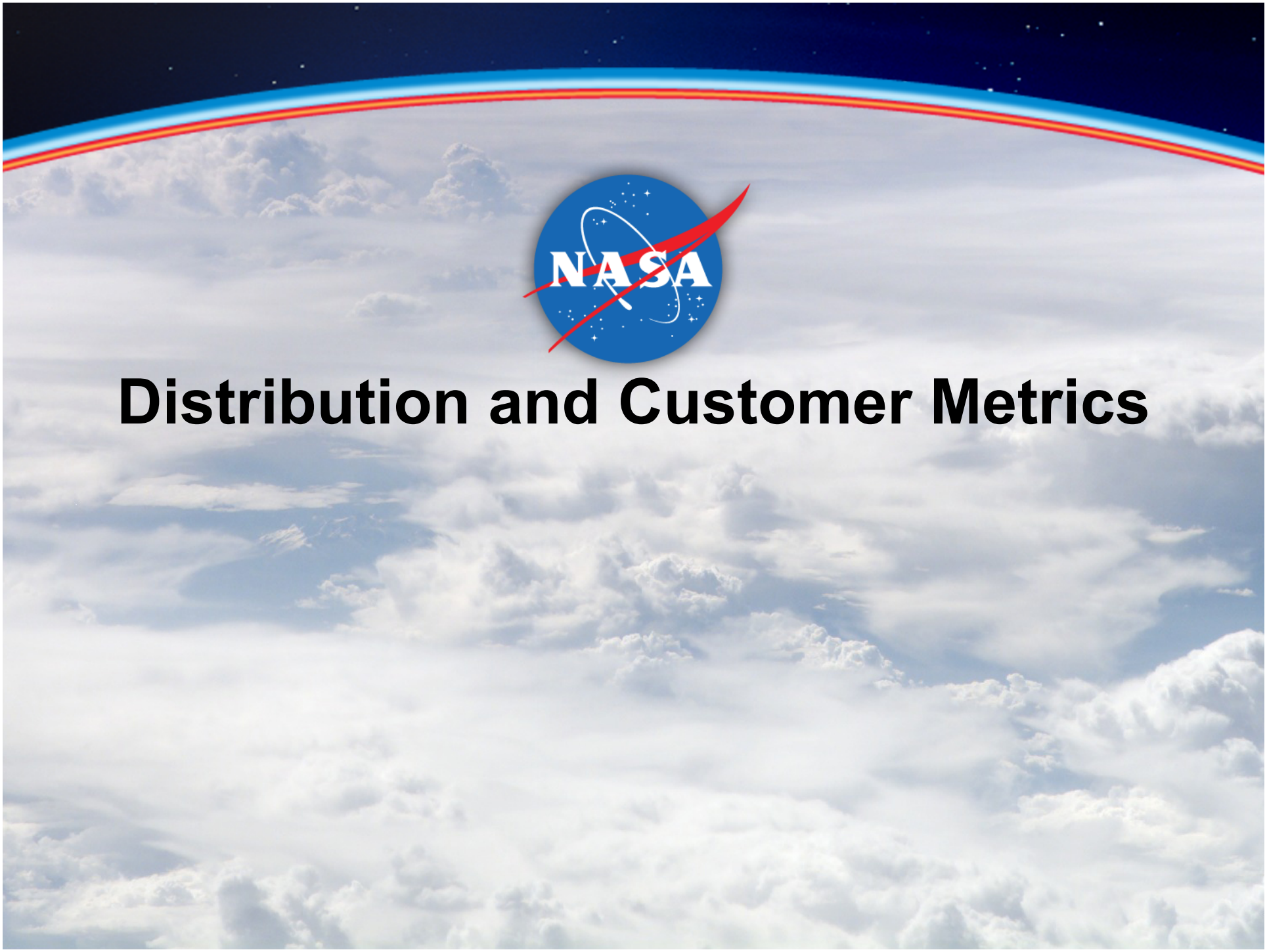
AMI-P


AMI-P Overview Architecture
2x SL8500 Tape Robots
Brocade DCX
IBMSAN 768B
DS4800
Disk Controller
GPFS Filesystem
Maximal configuration DS4800
12 Trays w/ 16 disks = 192 disks
38 LUNs @ (4 plus Parity) and 2 hot spares
8x 4Gb/sec fibre links
2x BladeCenter H
28 x Power6 Blades
112 Power6 Cores
3 x 3950 x86 Servers
48 AMD64 processing cores
2 x 3950 x86 Servers
Network Services Hosts
Network services hosts provide, local home directories, SGE master / shadow master, and NSD services.
1 x 3950 x86 Server
Data Motion Host
Dedicated host fibre connected to both switches for stage in and stage out of data. (Dedicated HBA for each switch)
Each Bladecenter: 8 x 8Gb/sec
Each 3950: 2 x 8Gb/sec
Total: 28 client connections.
Gigabit Ethernet
All AMI-P hosts should be on the same physical Ethernet switch.
This switch should be part of the AMI private network and does not need to be isolated.
All hosts can network mount AMI filesystems.
Interactive Hosts
SGE submission hosts network mount AMI-P filesystem.
Director Fibre Switches
AMI GPFS Filesystems

AMI DPO Data Population

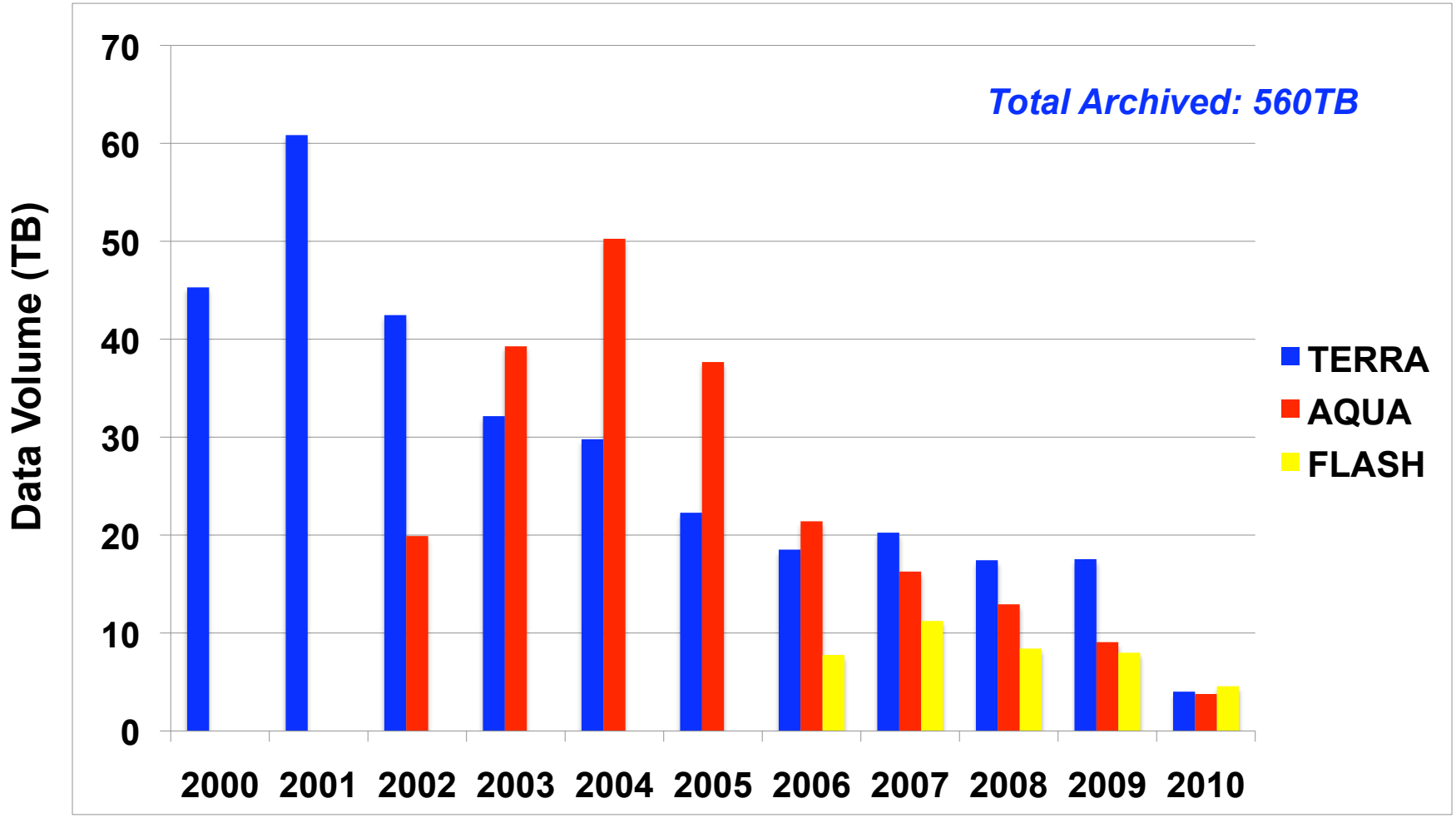
- Data products ingested and archived into ANGe and automatically populated to the AMI Data Products Online (DPO) Disk Cache:

Data Holdings in AMI DPO	As of Nov. 17, 2009		As of February 23, 2010		As of September 8, 2010	
Project Data Sets in DPO	**# files**	**Volume (TB)**	**# files**	**Volume (TB)**	**# files**	**Volume (TB)**
/ASDC_archive/**CALIPSO**	550,811	33	1,140,498	50	2,045,488	78
/ASDC_archive/**CERES**	**1,512,004**	**70**	**2,810,717**	**109**	**5,232,173**	**161**
/ASDC_archive/**CloudSat**					749	.2
/ASDC_archive/**FLASH**	**163,469**	**3.1**	**313,103**	**5.4**	**556,826**	**9.2**
/ASDC_archive/**GMAO**	**81,722**	**20**	**131,299**	**24**	**257,642**	**32**
/ASDC_archive/**ISCCP**	**131,575**	**.3**	**131,575**	**.3**	**198,548**	**.4**
/ASDC_archive/**MCIDAS**	**215,924**	**3.7**	**224,809**	**3.8**	**286,193**	**4.9**
/ASDC_archive/**MODIS**	**3,026,674**	**82**	**3,426,808**	**92**	**5,606,622**	**131**
/ASDC_archive/**NCEP**	**0**	**0**	**19,419**	**.2**	**54,567**	**.5**
/ASDC_archive/**SRB**	12,810	1.1	12,810	1.1	87,117	3.9
TOTALS	**5,699,276**	**214**	**8,211,038**	**286**	**11,628,072**	**421**

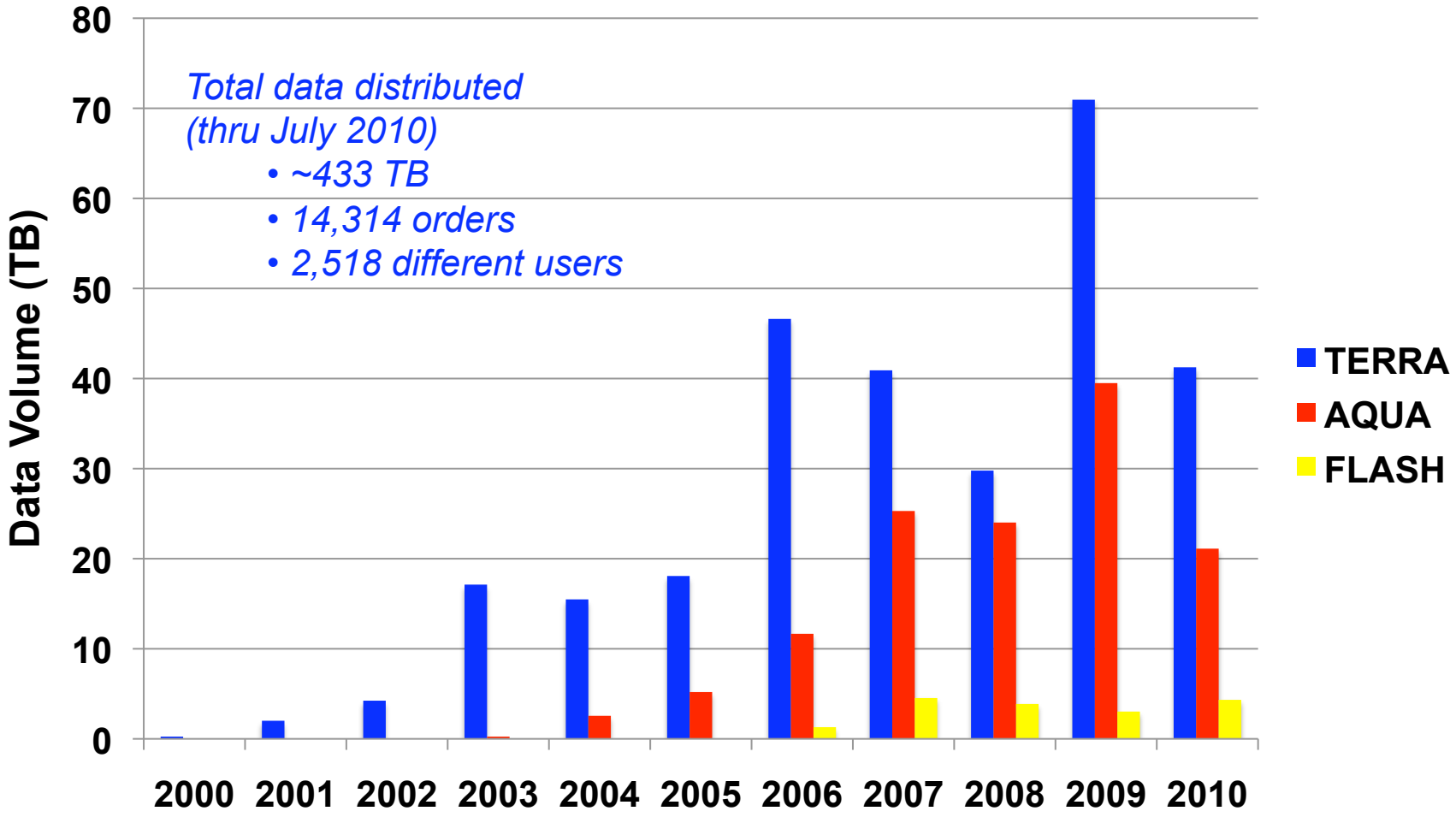


Distribution and Customer Metrics

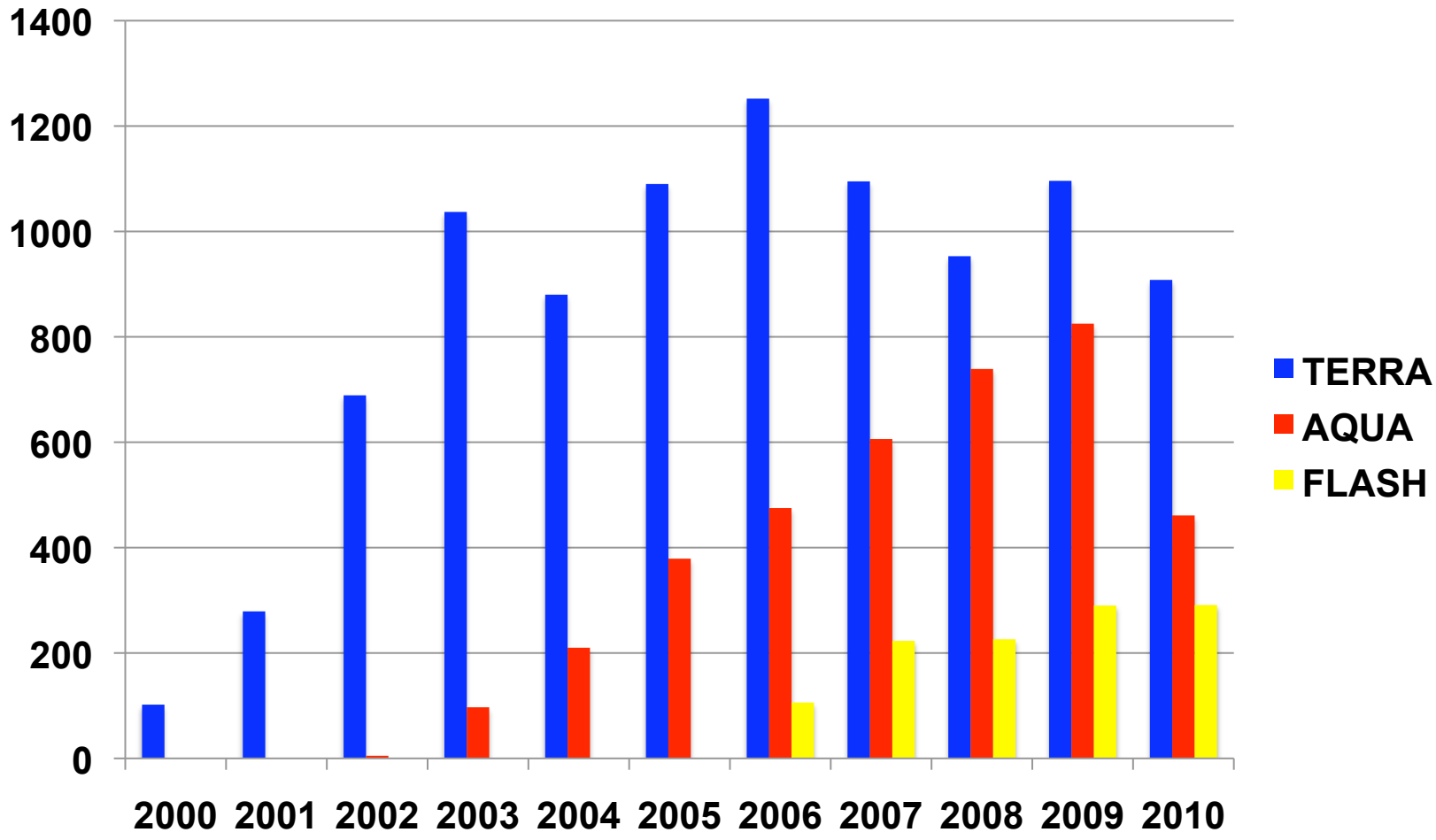
CERES and FLASHFlux Archive Volume by Data Year



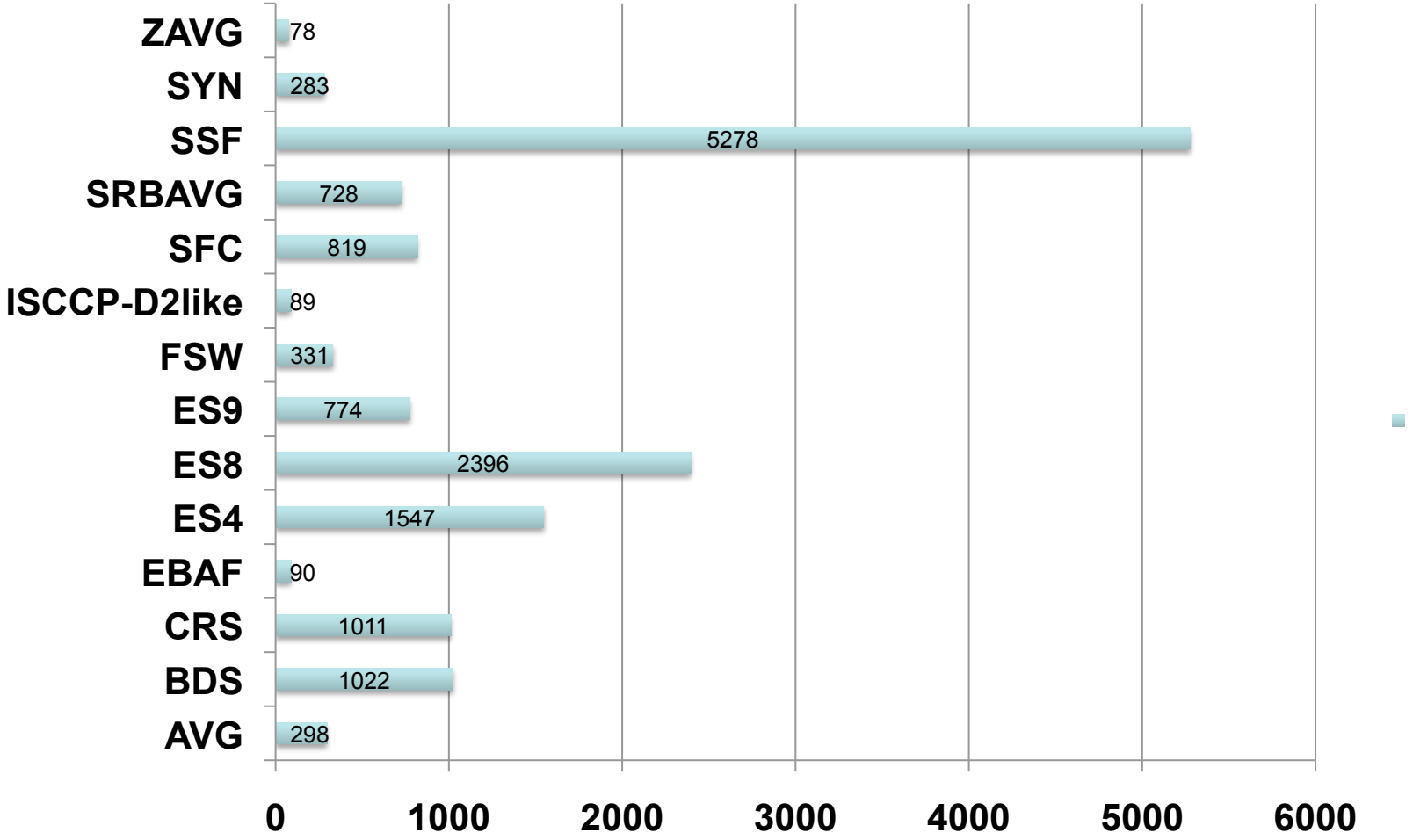
CERES and FLASHFlux Data Distribution



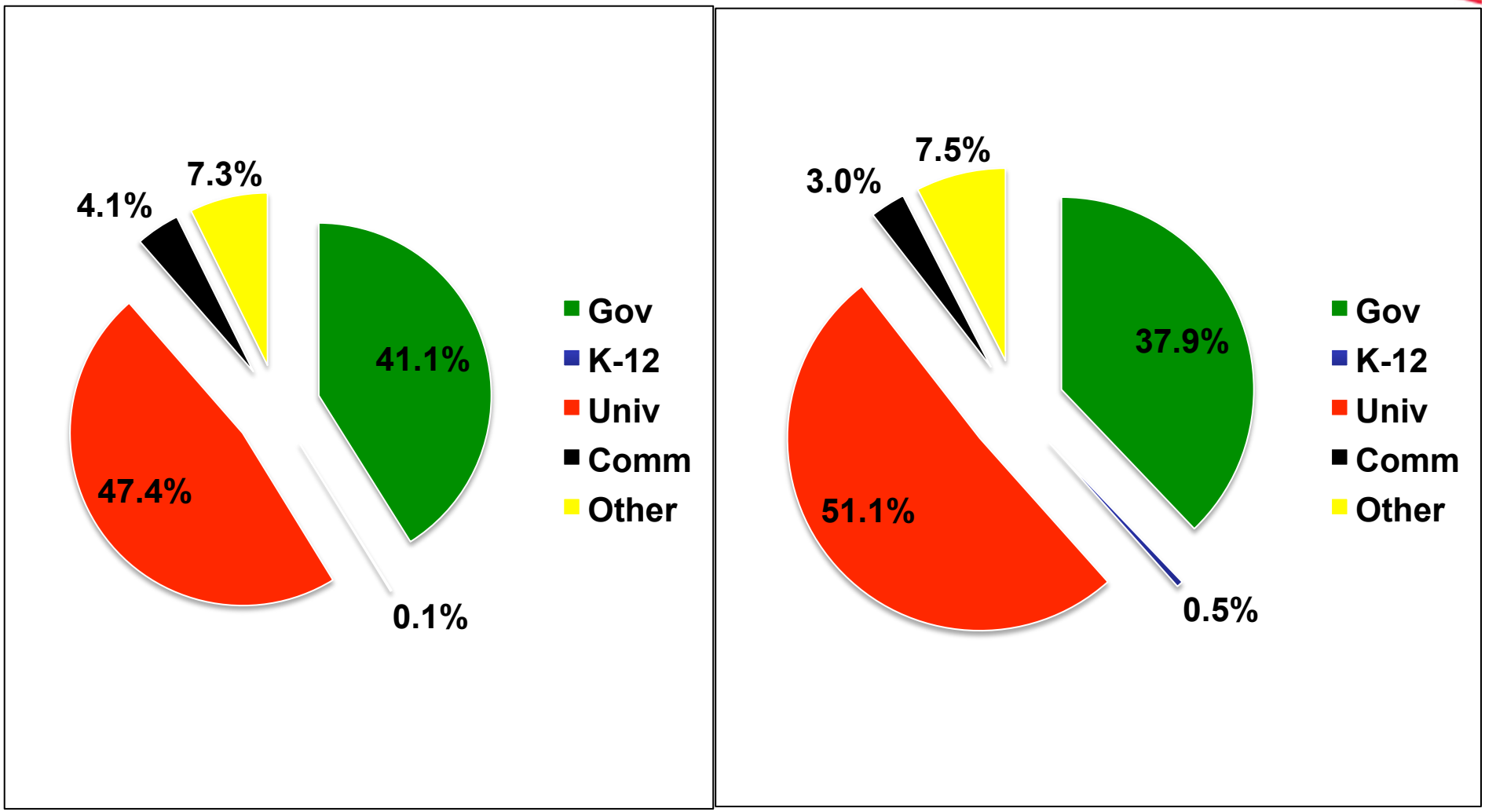
CERES and FLASHFlux Data Orders



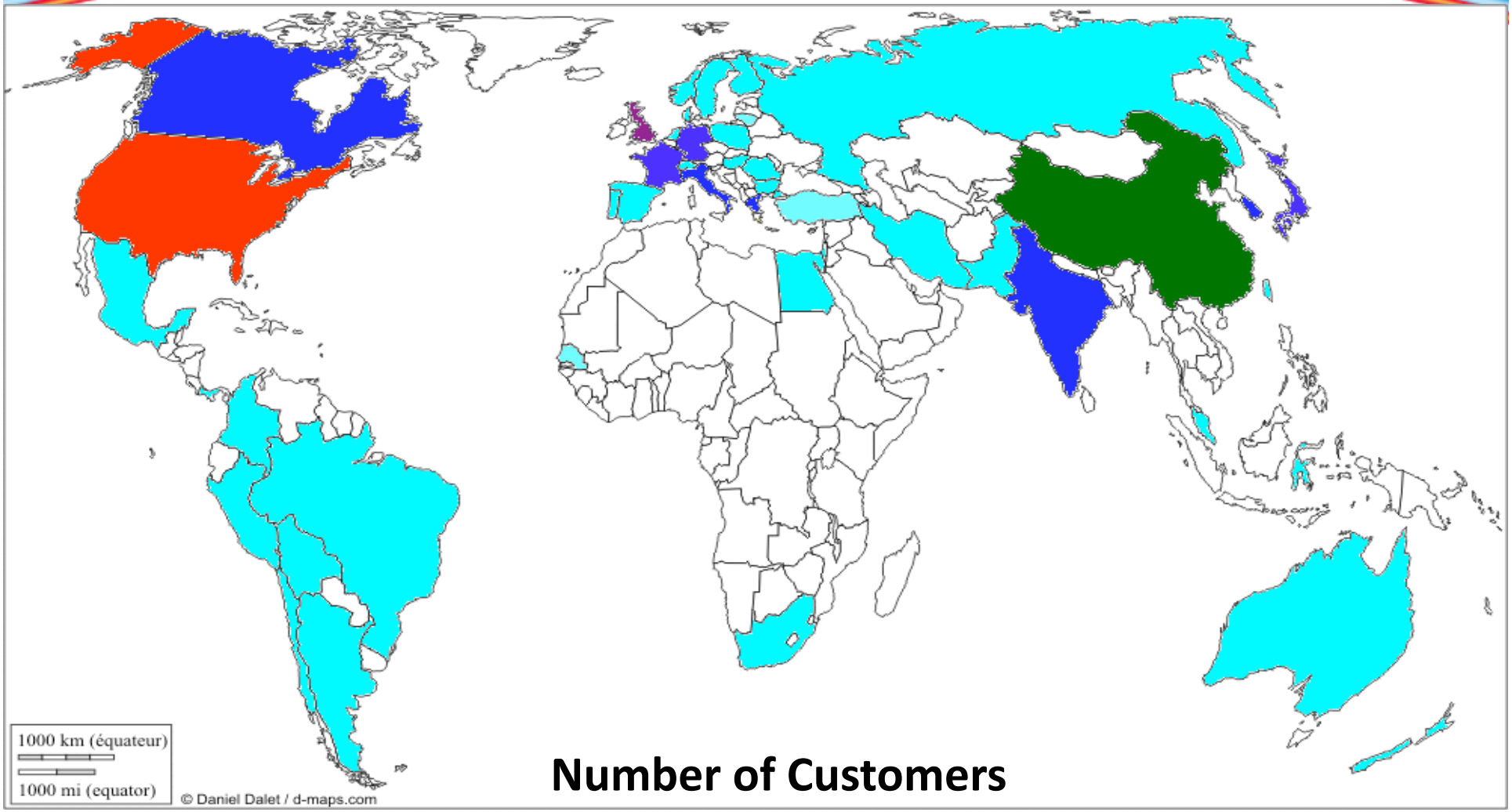
CERES Orders by Product (Mar 2000- Jul 2010)



CERES Customers by Affiliation



ASDC CERES Data Users



1 to 10 | 25 to 40 | ~60 | ~130 | ~890

Accessing CERES Data

- ASDC Web Page
  http://shire.larc.nasa.gov/HBDOCS/langley_web_tool.html

- ECHO (EOS ClearingHOuse) WIST (Warehouse Inventory Tool)
  https://wist.echo.nasa.gov/~wist/api/imswelcome/

- CERES Subsetter
  http://ceres.larc.nasa.gov/order_data.php

What's New at the ASDC

- CALIPSO Search and Subset Tool

http://www-calipso.larc.nasa.gov/search/

Future data sets supported:

- MISR
- TES
- CERES

Contacting ASDC User Services

User Services: larc@eos.nasa.gov

Web site: http://eosweb.larc.nasa.gov